

UNCLASSIFIED

Mathematical capture of human data for computer model building and validation

*Gladstone Reid, MSBME; Gordon Cooke, MEME; Robert Demarco, MSBME;
Caitlin Weaver, MSBME, MSME; John Riedener, MSSE.; Elizabeth Mezzacappa, PhD*

U.S. Army - Target Behavioral Response Laboratory
Research, Development and Engineering Command
Armament Research, Development and Engineering Command
Quality Engineering and System Assurance Sciences Division
Picatinny Arsenal, NJ 07806
973-724-9175
gladstone.v.reid.civ@mail.mil

Keywords:

Motion Capture, Crowd Management, Modeling, Simulation, Regression,
Validation, and Non-Lethal Weapons

ABSTRACT: *The model development process created at the TBRL has successfully created models from LVC simulation data, simulated models over time to generate simulated crowd data per condition, performed analysis of observed and simulated data, and compared simulated and observed data against crowd behavioral measures. The process allows for quantitative means for validation of both the mathematical (statistical) and computational models against empirical data. These findings support the claim that more accurate human behavioral models may be derived using laboratory observed data. Furthermore the findings support the necessity of a continuation of the work to explore ways to improve the model development and validation processes.*

1. Introduction

Analyzing the behavior of crowds has been of interest to the Department of Defense for years (Loftin 2005, McKenzie 2008). With the rise of urban warfare, an understanding of crowd behavior has become even more urgent, in particular to help decrease civilian casualty during rioting events and protests. Past and current civil unrest have shifted from peaceful to violent in moments, leaving military and law enforcement personnel with little time to strategize safe de-escalation. Understanding a crowd's dynamics, decision making, leadership, momentum, division of roles and motivation sources are essential to developing technology to deter, deny, suppress, disperse and de-escalate a crowd. This understanding is not gained by merely reviewing archived reports and videos of crowds in volatile states, but is gained through scientific analysis of the components that predict crowd behavioral response.

Live Virtual Constructive (LVC) simulations can help to build the understanding of crowds under controlled but realistic environments. LVC simulations allow for crowds to be immersed in operationally relevant scenarios, where each aspect that contributes to their behavior can be analyzed to determine the internal mechanisms that influence decision making. Non-Lethal Weapons (NLW) have been the primary force option during civil unrest. NLW manufacturers are challenged to design their products to be effective in

managing crowds while providing sufficient safety for its users and targets. LVC simulations are ideal venues to test NLW to evaluate their effectiveness in crowd management situations and to determine crowd responses to technology developed to mitigate their volatile effects.

The Target Behavioral Response Laboratory (TBRL) has deployed numerous LVC Simulations over the past eight years that have led to the determination of the effectiveness of several NLW energies and technologies. (Mezzacappa, E. S; Cooke, G. et al (2008); Short K. et al (2010)). TBRL researchers have leveraged data generated from LVC simulations to generate computational models to predict human behavior in tactically relevant scenarios; bridging the gap between laboratory data and modeling. This approach to modeling and simulation uses behavior of real persons as the analytical link to modeling and simulation. TBRL researchers have created a process using MATLAB software to develop and validate models for simulating human behavior with the use of data gathered during crowd management LVC simulations. The process discussed in this paper was developed to evaluate the effectiveness of crowd control weapon systems, but can be used in the evaluation of other systems. It is comprised of several modules that work together to produce simulated data for crowd locomotive behavior; estimating crowd responses to several non-lethal technologies and their surrogates.

Report Documentation Page			Form Approved OMB No. 0704-0188		
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE 03 APR 2014		2. REPORT TYPE Conference Proceedings Paper		3. DATES COVERED 00-00-2011 to 00-00-2014	
4. TITLE AND SUBTITLE Mathematical capture of human data for computer model building and validation Excerpted from the Proceedings of the 23rd Annual Behavior Representation in Modeling & Simulation (BRIMS) Conference, April 1-4, 2014, Washington, DC			5a. CONTRACT NUMBER		
			5b. GRANT NUMBER		
			5c. PROGRAM ELEMENT NUMBER		
6. AUTHOR(S) Gladstone Reid; Gordon Cooke; Robert DeMarco; Caitlin Weaver; John Riedener, et al			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Army, ARDEC, Target Behavioral Response Laboratory, RDAR-EIQ-SD, Building 3518, Picatinny Arsenal, NJ, 07806-5000			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)			10. SPONSOR/MONITOR'S ACRONYM(S)		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S)		
12. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution unlimited					
13. SUPPLEMENTARY NOTES Last author is Elizabeth Mezzacappa.					
14. ABSTRACT The model development process created at the TBRL has successfully created models from LVC simulation data, simulated models over time to generate simulated crowd data per condition, performed analysis of observed and simulated data, and compared simulated and observed data against crowd behavioral measures. The process allows for quantitative means for validation of both the mathematical (statistical) and computational models against empirical data. These findings support the claim that more accurate human behavioral models may be derived using laboratory observed data. Furthermore the findings support the necessity of a continuation of the work to explore ways to improve the model development and validation processes.					
15. SUBJECT TERMS Motion Capture, Crowd Management, Modeling, Simulation, Regression, Validation, and Non-Lethal Weapons, Target Behavioral Response Laboratory					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT Public Release	18. NUMBER OF PAGES 8	19a. NAME OF RESPONSIBLE PERSON
a. REPORT unclassified	b. ABSTRACT unclassified	c. THIS PAGE unclassified			

2. Background

Human behavioral science studies human actions and seeks to generalize behavior in society. LVC simulations have a direct link to behavioral science in that it facilitates the studies necessary to gain the understanding to generalize human behavior. Like most research efforts there is a hypothesis at the onset that drives the testing, however the analysis of the data leads the investigator to the actual finding. This is the concept behind TBRL's LVC simulations: LVC simulations are deployed to investigate the behavioral response of individuals when faced with NLW to determine quantitative measures of effectiveness. The data from these simulations can be used to understand the general principles behind human behavior and serve as the input for building models that represent human behavior. With accurate validated models, virtual simulations can be developed to predict human behavior that will fill some of the military's need to "understand the motivations and influences underlying adversarial behavior, behavior of contested populations, and populations with whom US Forces have not yet interacted, how they vary cross-culturally, and what is innately human behavior that extends across cultural boundaries is required at all levels of military operations." (Office of Naval Research, 2008).

The TBRL has developed a scientific approach to solving the military's needs to understand crowd behavior. The conceptual model for the TBRL crowd research program is built on Lewinian Field Theory, which proposes that human behavior can be explained as attractions and repulsions toward and away from goals (Lewin, 1935). The TBRL has used their Crowd Behavioral Test-Bed to gather locomotive, psychosocial, and effectiveness data from several groups of crowds to develop models that use vector regression methods to identify attributes of a crowd that influence predictive variables. The steps can be summarized as collecting crowd behavioral data, processing the data to serve as an input to a model building algorithm, computing a mathematical (statistical) model, computing simulations of crowd locomotive behavior, processing both experimental and simulated data to gather crowd metrics, and comparing experimental and simulated crowd metrics to determine validity.

3. Methods

3.1 Crowd Behavioral Data Collection

TBRL's Crowd LVC simulation was deemed as human subject research by the ARDEC Human Subjects Protection Administrator and was vetted through an Institutional Review Board (ARDEC IRB# 10-0002

Effectiveness Testing for Crowd Management). The simulation involved the configuration of an indoor test bed (See Figure 3.1.1), recruitment of human subjects, collection of data, processing and analysis of data, and reporting of results and conclusions.

The primary elements of the test-bed were developed and verified during the implementation of the Crowd Behavior Test bed (CBT) that incorporated Vicon® Motion Capture (MoCap) hardware and software, digital video cameras (Cooke et al, 2010). Simulated crowds with up to 25 subjects during a trial, wore uniquely configured helmets to allow for capture of their location with six degrees of freedom. In addition to these elements there was a need to construct a linear goal system as shown in Figure 3.1.1, install a Medium Range Acoustic Device (MRAD), develop a simulated Area Denial Technology (ADT) non lethal weapon output, and develop custom LabVIEW software for data collection and device triggering.

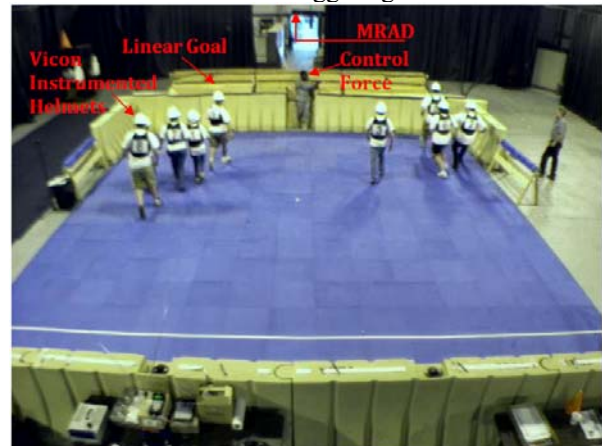


Figure 3.1.1: Crowd Behavior Testbed

Human subjects were recruited from the local and surrounding communities via flyers, and communication with previous volunteers. Subjects were paid \$20 per hour of participation and had the ability to earn more during participation. Experimental conditions consisted of the subjects seeking an opportunity to successfully get their simulated rock into the linear target while the target was defended by a non-lethal device. In two of the five conditions the target is defended by a Control Force (CF) that uses a non-lethal device. In both scenarios the CF defends the target with a simulated hand-held standoff non-lethal device. One of these devices represents a projectile weapon and the other a visible directed energy device. For two other conditions, the target is being defended by either an Medium Range Acoustic Device (MRAD), or a simulated invisible directed energy (IDE) weapon. The Projectile, the VDE, and the IDE weapons had effects of financial loss for the targeted participant, while the MRAD yielded its own penalty through loud sound exposure. No-weapon conditions were also run to provide baseline comparison data.

3.2 Model Building

Data from LVC simulations are useful within themselves, however the TBRL has constructed a novel process to re-use this data to build and validate models, which in turn can create additional simulated data (Patent Submission Docket 2011-047). This simulated data is very useful for instances where LVC simulations are not feasible or possible. The method of building models from LVC simulations gives more validity and confidence to the simulated data, since it is generated from real data and not subject matter expert opinion. The goal of creating a model building and validation process yielded the requirements: to create software to automate the creation of mathematical models for motion of individuals in a crowd from the data collected in TBRL crowd experimentation, to create software to allow running simple simulation of the above model to generate simulation data, to create software for the calculation of crowd level metrics from laboratory and simulated data, and to create software to compare simulated data with observed data. The block diagram in Figure 3.2.1 outlines the modules created in MATLAB to satisfy the requirements.

3.3 Pre Process Module

The pre-process module processes the raw motion capture data into a form that can be used in subsequent processes to derive the mathematical model. Vicon MoCap system exports location data in comma separated value format, storing all subject data for each trial in one file where each subject's data are appended to the end of the preceding subject. For the purpose of modeling, there is a need to have a standard format for model input where data are organized in a manner so that the dependent and independent variables are easily distinguished for each condition. This requirement led to the development of the pre-process module in MATLAB to restructure the location data file into an ASCII text file where each row within a two-dimensional matrix represents data for one subject for a time step and is followed by additional data for other subjects for that same time step and is followed in the same order for additional time steps.

3.4 Input Module

The input module parses the formatted data from the previous step into three elements: a header vector, an output matrix and a predictor matrix. The header vector is derived from the first row of the data file and includes the headers of all the columns that are included in the predictor matrix. The output matrix ($n \times 2$) consists of the 'Vx' and 'Vy' velocity vector values ('n' represents the number of data points). The predictor matrix ($n \times p$) consists of all data included in the data file excluding the output matrix and the header vector ('p' represents arbitrary number of predictor variables). The predictor matrix is formatted, where Columns 1 through 5 are designated to subject identification number, time elapsed (seconds), 'X' location relative to CF, 'Y' location relative to CF, and run identification number respectively.

3.5 Modeling Module

The modeling module accepts as input, the predictor and output matrices to create the mathematical equations reflecting the relationship between predictor/input and predicted/output variables using non-linear regression. There are two testing paradigms in the Crowd Management experiment; a condition that only uses the target (baseline) with no protection by non-lethal devices and another where the target is protected by either a control force member and/or a non-lethal device. To accurately model the behavior of the person, their behavior from the single influence of the target alone (baseline) was used to generate a model of the attractive force generated by the target. This model was used to subtract the force created by the target out from the data collected when both the target and a non-lethal device were present; the result of the subtraction provides values for the influence of the non-lethal device alone.

The module uses the non-linear fit function from the (Mathworks, Inc. Nonlinear Regression, 2011) Statistical Toolbox to perform the non-linear regression. The syntax for the non-linear fit MATLAB

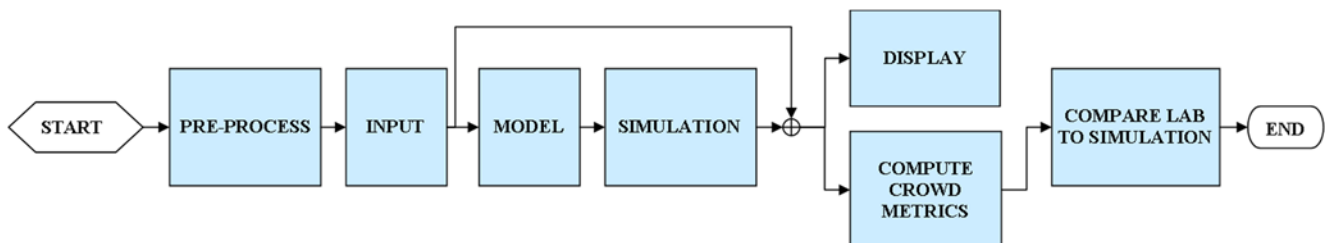


Figure 3.2.1: TBRL Model, Simulation and Validation Process

function is `[beta, r, J, COVB, mse] = nlinfit(X, y, fun, beta0)`, where 'beta' are the fitted coefficients, 'r' are the residuals, 'J' is the Jacobian of the function (fun), COVB is the estimated covariance matrix for the fitted coefficients, and 'mse' is an estimate of the variance of the error term. In the syntax 'X' stands for the predictor variables, 'y' represents the responses/output, 'fun' represents a model function and 'beta0' represents the starting coefficients.

The module does the regression of velocity vectors in both 'X' and 'Y' axes against the predictors, generating model coefficients for change in location in 'X' and 'Y' coordinates, along with confidence intervals for input values. The model, the predictors, and responses are then processed to determine the model errors that are then fit to the Weibull distribution. The module is designed to incorporate the effects of the CF differed from the above description in that it also converts the coordinates from Cartesian to polar coordinates with the origin at the CF location. The polar coordinates are used to fit the model instead of the Cartesian coordinates.

3.6 Simulation Module

The simulation module is built to execute a time stepped simulation of each subject's behavior based on the derived model, start conditions, average time between samples, and duration of simulation. At each time step the change in distance traveled, the change in position, and velocities are calculated for each subject and concatenated to a simulated data file, structured in the same manner as the predictor matrix. In addition to the basic simulation module that simulates the baseline conditions, there is one that is designed to incorporate CF effects by transforming the coordinates of the baseline model to fit that of the CF model, where the CF location is the origin for polar coordinates.

3.7 Crowd Metric Module

The crowd metric module calculates the leading edge (LE), trailing edge (TE), centroid, geometric center, and dispersion for the crowd (Cooke et al, 2010). These measures are considered aggregate metrics of crowd behavior as a whole, rather than individual paths taken by each subject. The LE is defined as the location of the forward most crowd member for each experimental trial with respect to the target or CF for each data time-step. The TE is defined as the location of the crowd member that is furthest back for each experimental trial with respect to the target or CF for each data time-step. The centroid is the location of the crowd member that is at the midpoint between the LE and TE during each time-step. The measures are

calculated using a sorted matrix, 'y_data', in which rows represent each time step and columns represent individual subjects. The function calculates the LE by finding the maximum 'y_data' point for each run. The TE is calculated similarly but with the minimum 'y_data' point. The centroid is calculated by taking the mean of the 'y_data'. The geometric center is calculated by taking the average of the leading and trailing edge. Dispersion is calculated by evaluating the average displacement in the 'X' and 'Y' direction. The function then plots the LE and centroid for visual verification.

3.8 Model Comparison

The model comparison module compares, statistically, the output of the simulation with the observed human data collected in the laboratory. The observed data used in the comparison is a set collected under the same conditions as the data used for modeling. For example, an available data set can be split into data for modeling and data for comparison with simulation output. This process is akin to establishing "split-half reliability" in behavioral science. This module uses two sample Kolmogorov-Smirnov (K-S) goodness of fit (GOF) test (`kstest2`) to determine goodness of fit for crowd metrics (Mathworks, 2011). The function accepts two sets of data, significance level (α), and type of alternative hypothesis test. We used the `kstest2` function to compare the cumulative distribution functions (CDF) of the observed and simulation data to determine: if the simulated data follows the same distribution as the observed data (hypothesis acceptance), the asymptotic p-value, and k-statistic. The null hypothesis is that both observed and simulation data are from the same continuous distribution. If the null hypothesis (h) is accepted, 'h' value is 0, but 1 if rejected. The k-statistic is the greatest distance between the CDF plots of the observed and simulation. For the model comparison module 5 % significance level was used and the default of unequal alternative hypothesis test.

4. Results and Discussion

In this section we discuss the results of the comparisons between 1) the output of the simulation based on the mathematical model developed from the empirical data and 2) the human behavior data collected in the lab. A successful comparison is indicated by derivation of quantitative metrics of the goodness of fit between the output and the data. The efforts in this project developed methods to quantify improvements to the model, that is, quantitative metrics of model validation.

Both the mathematical model (the equation derived from the data) and computational model (the equation

used in the simulation) were compared to the data collected in the laboratory. The mathematical models generated through the process were evaluated based on root mean squared error (RMSE) and graphical comparison. The computational model was evaluated by the K-S GOF test.

4.1 Mathematical Model Validation: RMSE

This exercise was done using individual locations at time-points during the experimentation. A mathematical model predicting current location of a crowd member based on previous location, and velocity was derived using regression methods. The RMSE is a comparison between the expected values calculated from the regression equation and the data observed in the laboratory. The RMSE derived for the models of the LVC simulations can be found in Table 4.1.1; the larger the RMSE, the greater the discrepancy between the expected values and observed values of behavioral measures in the laboratory. As seen in Table 4.1.1, the largest error is with the radial model for the projectile weapon, where the error is equivalent to approximately 2 ft. This error is acceptable when we consider most individuals being modeled are approximately 2 ft wide. The RMSE for all models are below 0.66 meters, which would generally mean the model is a good representation of the input data, however the graphical comparison shows that there is a discrepancy when the model is run through simulations.

Table 4.1.1: Model mean standard error values

Conditions	RMSE (Radial) [meters]	RMSE (Tangential) [meters]
Baseline	0.5	
Projectile Weapon	0.65	0.54
Visible Directed Energy	0.47	0.36
MRAD	0.37	0.26
Invisible Directed Energy	0.37	0.38

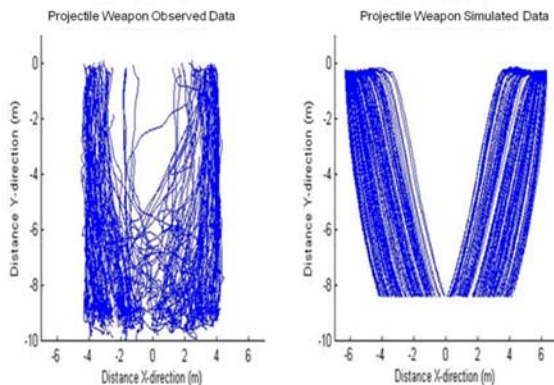


Figure 4.1.1: Baseline Observed vs. Simulation

4.2 Graphical Comparisons

Graphical comparison involved graphing the experimental trials for each condition adjacent to the simulation results derived from applying the model over numerous iterations for each respective condition. This provided a quick examination whether or not the model generated simulation results were visually

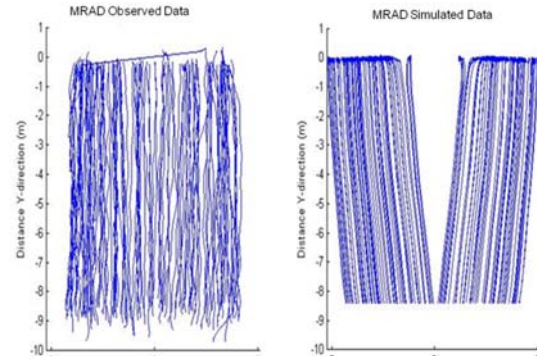


Figure 4.1.2: MRAD Observed vs. Simulation

similar to the observed movement patterns. The graphical comparison of the observed and simulated data for the models with largest and smallest RMSE from Table 4.1.1, the projectile weapon and MRAD respectively are graphed in Figures 4.1.1 and 4.1.2.

The graphical comparison of the projectile weapon, in Figure 4.1.1 shows a very similar representation of the observed data. The comparison reveals that the model does not account for some situations where crowd members take a direct or a diagonal approach. Since these were the minority, it would be expected that the model may treat these data as outliers during model regression. The RMSE associated with the projectile weapon is the largest of the five conditions, even though the graphical comparison may be judged to show the best visual representation.

The MRAD comparison in Figure 4.1.2 shows that the simulation trends to the extreme of having crowd member approach from either the right or left of the stimulus which is not representative of the observed data, which shows some crowd member taking a direct approach. With this comparison it can be seen that even though the RMSE indicates the best fit, the model does not account for the common behavior of taking a direct approach from the center. This effect may be due to the density of the paths on the sides versus the density in the middle and how those data points were weighted in the model generation.

4.3 Computational Model Validation

The computational model used to generate the simulations consisted of the regression equation, that is, the mathematical model, plus a stochastic component derived from the data. This stochastic component consisted of a randomly generated error

term that was selected from a distribution of the error between the observed and predicted data for all crowd members. The validity of the computational model, was tested using the K-S GOF test on aggregate crowd measures (LE, TE, and Centroid) comparing the CDF of forecasted behaviors to the CDF of observed behaviors.

The K-S GOF was performed using crowd aggregate measures to determine if the simulated crowd produced similar crowd aggregate locomotive behavior to the observed. The K-S Statistic, which represents the largest vertical distance between the observed and the simulated CDF plots, is a measure of how closely the simulated fits to the observed data. The mean crowd aggregate measures: LE, TE, and centroid of the simulation data were compared to the observed data at an α of 0.05. The results for the K-S GOF are shown in Table 4.3.1 and Figures 4.3.1, 4.3.2, and 4.3.3.

The K-S GOF test failed to reject the null hypothesis for LE, Centroid and TE for the baseline and The VDE weapon conditions since p-values exceeded α . All other conditions rejected the null hypothesis except the LE for VDE weapon. The K-S Statistics were correspondingly lower for the measures that failed to reject the null hypothesis. The CDF plots for Baseline and MRAD conditions in Figures 4.3.1 and 4.3.2 show a very similar locomotive behavioral response to the stimulus by the simulated crowd compared to the observed.

Table 4.3.1: K-S Goodness of fit results

Conditions	Measures	H [Accept(0) /Reject(1)]	P-value	K-S Statistic
Baseline	LE	0	0.188	0.500
	Centroid	0	0.519	0.375
	TE	0	0.519	0.375
MRAD	LE	0	0.883	0.286
	Centroid	0	0.883	0.286
	TE	0	0.883	0.286
Projectile Weapon	LE	1	0.001	0.714
	Centroid	1	1.19E-04	0.786
	TE	1	1.87E-05	0.857
Visible Directed Energy	LE	0	0.077	0.667
	Centroid	1	1.19E-04	0.786
	TE	1	1.87E-05	0.857
Invisible Directed Energy	LE	1	1.87E-05	0.857
	Centroid	1	2.50E-06	0.929
	TE	1	2.86E-07	1.000

For the projectile weapon it can be seen from Figure 4.3.3 that the model estimated the crowd would approach the goal location, however, the actual crowd chose to stay back and try to achieve their goal from a distance (they threw from farther away, rather than approaching closer to the target). This is indicative that the model does not accurately predict the locomotive behavior of the crowd during this condition. This

reveals a possible problem in the weighting of location data in the model development process. If the model is solely developed from the regression of location data, it will have limitations as seen in these conditions. Since crowd members did travel to the goal location for each condition, the regression analysis considered those locations, however from the CDF plots it can be shown that on average the crowd approaches the goal location more cautiously and sometimes hangs back for the entire trial. If the LE, TE, and centroid are incorporated in the model building process it may account for these crowd behaviors that are not replicable through regression analysis.

All methods of analyzing the simulated data proved useful, the RMSE revealed the estimation error of the mathematical model. The graphical comparison of the observed and the simulation data revealed that even the small error in estimation shown in Table 4.1.1, can drive the simulated data outside the bounds of the observed behavior. The K-S GOF test identified aspects of the computational model of crowd behavior

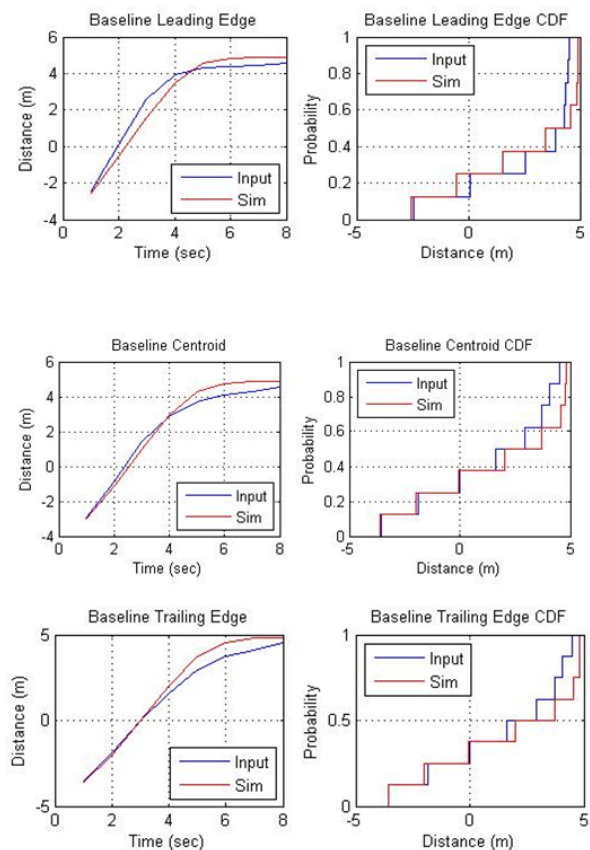


Figure 4.3.1: Baseline Crowd Measures K-S GOF comparisons (Leading Edge, Centroid, and Trailing Edge)

that may need to be included to yield a correct representation of how the crowd behaves in response to various stimuli. All methods highlight limitations to the models generated, but show promise that the process

developed is capable of accurately predicting crowd behavior in response to stimuli.

5. Summary and Conclusion

The model development process created at the TBRL has successfully created models from LVC simulation

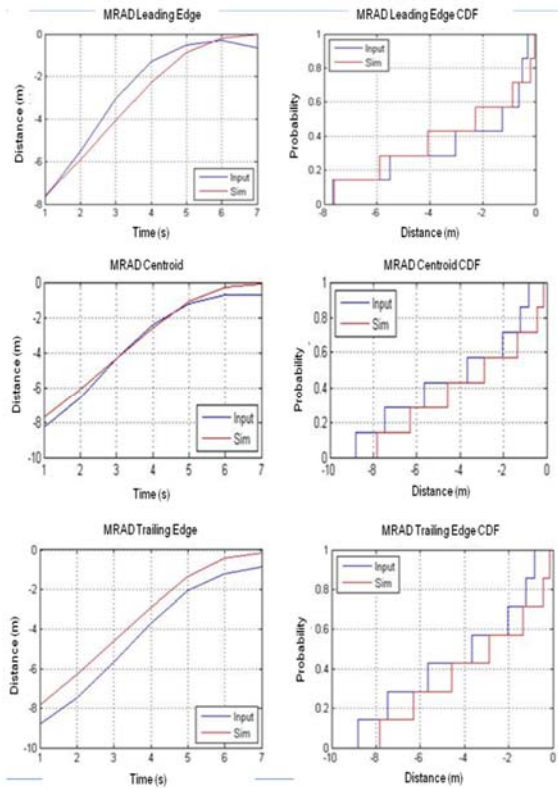


Figure 4.3.2: MRAD Crowd Measures K-S GOF comparisons

data, simulated models over time to generate simulated crowd data per condition, perform analysis of observed and simulated data, and compared simulated and observed data against crowd behavioral measures. The process allows for quantitative means for validation of both the mathematical and computational models against empirical data, such as in the following example.

The models created for the baseline and The VDE weapon conditions successfully estimated crowd locomotive behavior based on their RMSE values of 0.5 and 0.37 meters respectively and further validated by the K-S GOF test. The K-S GOF test determined that the simulated data generated from the models of the baseline and MRAD are from the same distribution as that of the observed data at an α of 0.05. The K-S statistic for the crowd measures LE, TE, and centroid for the baseline condition are 0.5, 0.375, and 0.375 respectively. The K-S statistic for the crowd measures

for the VDE weapon condition is 0.286 for all measures.

The models created for a projectile weapon, the simulated VDE weapon, and the simulated IDE weapons still need refinement based on the graphical comparison of the simulated and observed data in addition to the K-S GOF test. The RMSE values were

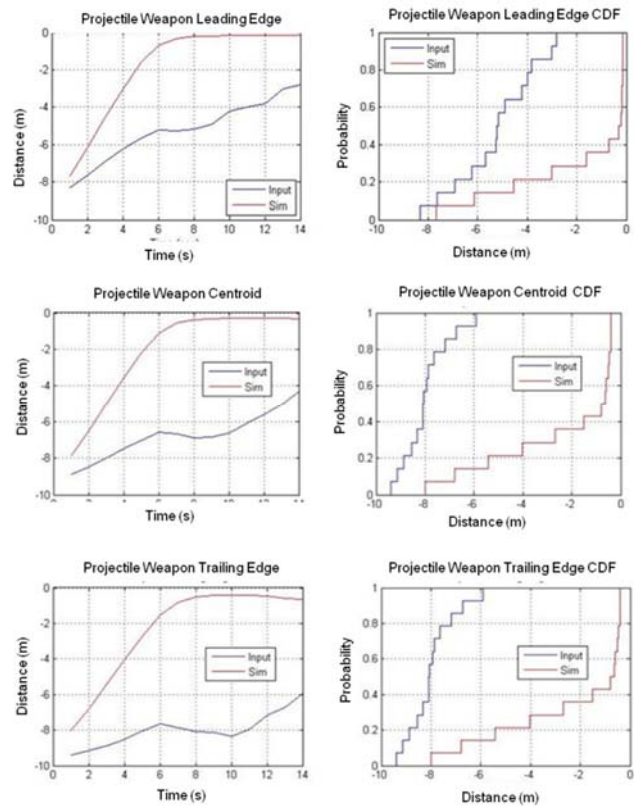


Figure 4.3.3: Projectile Weapon Crowd Measures K-S GOF comparisons

favorable, and the graphical comparison for the projectile weapon and VDE weapon were relatively precise, however the K-S GOF test revealed that the mean LE, TE, and centroid of the simulated data were not from the same distribution as the observed data for those conditions at an α of 0.05 with p-values significantly lower than α .

There are several identified areas for improvement which include, 1) using additional or new regression methods that yield more accurate models reflected by smaller RMSE; 2) Accounting for error during simulation and making necessary adjustments to produce data that more closely represent observed data; 3) The incorporation of crowd metrics for the observed data set into the model development process to improve the goodness of fit for the observed and simulated data; and 4) the use of additional goodness of fit test to validate the model.

In conclusion, these findings support the claim that more accurate human behavioral models may be derived using laboratory observed data. Furthermore the findings support the necessity for a continuation of the work to explore ways to improve the model development and validation processes to derive a robust model that can be used to simulate crowd behavior for the purpose of evaluating effectiveness of weapons systems used to defend against hostile crowds.

6. References

- Cooke, G.; Mezzacappa, E.; Sheridan, C.; DeMarco, R.; Tevis, K.; Reid, G.; Jaffery, N.; Riedener, (2010) Topology and individual location of crowds as a measures of effectiveness for non-lethal weapons. Proceedings of the 27th Army Science Conference, 29 November –2 December 2010.
- Cooke, G.; Mezzacappa, E.; Sheridan, C., Riedener, J., Yagrish, K.; (2008). Unmanned Aerial Vehicle Mounted Light Based Deterrence Effectiveness Experimentation (TBRL-ULBD-TR-08012).
- Kolmogorov-Sminov test (2011). Two sample K-S test (kstest2), Statistical Toolbox, MATLAB.
- Lewin, Kurt, PhD (1935). A Dynamic Theory of Personality: Selected Papers. New York:McGraw Hill Book Company, Inc.
- Loftin, R. B., Petty, M. D., McKenzie, F. D., & Gaskins, R. C. (2005). Modeling Crowd Behavior for Military Simulation Applications. In K. R. W. B. Rouse, OrganizationalSimulation (pp. 471-536).
- McKenzie, F. D.; Petty, M. D.; Kruszewski, P. A.; Gaskins, R. C.; Nguyen, Q. H.; Seevinck, J.; et al. (2008). Integrating crowd-behavior modeling into military simulation using game technology. *Simulation & Gaming*, 39, 10-38.
- Mathworks, Inc. Nonlinear Regression (2011). Nonlinear Regression (nlinfit), Statistical Toolbox, MATLAB. Mathworks, Inc.
- Mezzacappa, E. S. (in press) Effectiveness testing of non-lethal weapons. *Journal of Defense Modeling and Simulation*.
- Office of Naval Research Broad Agency Announcements (2008). Human, Social, Cultural and Behavioral Sciences basic research, applied research and advanced technology development opportunities.
- Patent Application Submission: Docket 2011-047
- Short, K.R., Reid, G., DeMarco, RM., Sheridan, C., Cooke, G., Galonski, L., Hedderich, E., Minor, T.R.,and Riedener, J., 2010: Blunt impact as a deterrent under nominally low approach motivation conditions. Report to Joint Non-Lethal Weapons Directorate, TBRL-JAT-TR-10004, 1-33.

Author Biographies

GLADSTONE REID, MSBME, is a Project Lead Engineer with expertise in experimental automation design and implementation.

GORDON COOKE, MEME, is a Principal Investigator at the TBRL. He was also a Chief Engineer, USMA, 11th Armored Cavalry Regiment stationed at the National Training Center, Fort Irwin, CA and served a one-year tour in Mosul, Iraq as the Deputy Division Plans Officer for Multi-Nation Division North-West.

ROBERT DEMARCO, MSBME, is a Project Lead Engineer and Certified LabVIEW Associate Developer.

CAITLIN WEAVER, MSME, MSBME, is a Biomedical Engineering Graduate Student at the Center of Injury Biomechanics at the Wake Forest University School of Medicine.

JOHN RIEDENER, MSSE, is the Technical Director of the TBRL. He is a Certified Lean/Six Sigma Black belt with a focus on Design of Experiments and Statistical Analysis of Test Results. Before coming to the TBRL, Mr Riedener was the Performance Management Lead for the ARDEC Software Engineering Center and was integral in leading the that organization to an elusive CMMI Level 5 maturity rating.

ELIZABETH MEZZACAPPA, PhD, is a Principal Investigator at the TBRL with expertise in social psychology, stress, and human behavior modeling and simulation issues.